10. ročník Žilina 2. – 4. 7. 2018
Otvorený softvér vo vzdelávaní, výskume a v IT riešeniach
osscconf.soit.sk
OSS Conf

OSSConf 2026: 1–7

Mapa míst konferencí ROBUST

Ondřej Vencálek (CZ) a Pavel Stříž (CZ)

Abstrakt. Článek je úvodem do práce s mapami s dodatečnými popisky ve výpočetním prostředí R. Během naší práce používáme data z Eurostatu plus interní informace o konferencích ROBUST. Vytvořená mapa je zobrazena v R i přes Rscript, vyexportována do png a pdf určeném pro Informační bulletin na tisk, navíc zobrazena ve webovém prohlížeči offline i online za pomocí knihovny plotly a jejich prostředí Chart Studio určeném pro práci v týmech a zaměřeném na zveřejňování grafů.

Klíčová slova. programovací jazyk R, Eurostat, sf, dplyr, ggplot2, ggrepel, plotly, Chart Studio, tcltk.

A Map of ROBUST Conference Venues

Abstract. The article is a brief introduction to a map creation in R programming language with additional information. We primarily use data from the Eurostat servers and internal information about the ROBUST conference series. A preview of the map is shown in R and Rscript environments, exported to the png image and the pdf file for printing purposes of the Information Bulletin as well as previewed offline and online using the plotly library and their Chart Studio designed for collaborative work and chart publishing.

Keywords. R programming language, Eurostat, sf, dplyr, ggplot2, ggrepel, plotly, Chart Studio, tcltk.

1. Úvod

Konferenční řada ROBUST má v roce 2020 20+20. výročí, a proto jsme se rozhodli vytvořit tzv. „robust(n)í mapu“, tj. mapu Česko-Slovenska se zvýrazněním míst. V neformálním duchu jako jsou tyto konference, možná jako dárek s cílem si zkusit pár fint v R.

2. Data z Eurostatu

Z webové stránky `https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts` stáhneme si soubor zip v řádku NUTS 2016, 14/03/2019, 1:1 Million, sloupec SHP. NUTS neboli Nomenklatura územních statistických jednotek jsou územní celky vytvořené pro statistické účely Eurostatu. NUTS 0 odpovídá státu, NUTS 1 území (ČR), NUTS 2 regionům, NUTS 3 krajům, LAU 1 (dříve NUTS 4) okresům a LAU 2 (dříve NUTS 5) obcím.

Užijeme adresář, kde budeme mít zdrojový kód v R. Konkrétně:

```
$ wget https://ec.europa.eu/eurostat/cache/GISCO/distribution/\
> v2/nuts/download/ref-nuts-2016-01m.shp.zip
```

Ze staženého souboru potřebujeme extrahovat jen jediný. Na vykreslení mapy nám stačí.

```
$ unzip ref-nuts-2016-01m.shp.zip NUTS_RG_01M_2016_4326.shp.zip
```

RG značí regiony (angl. regions; vedle toho najdeme BN/boundaries pro hranice a LB/labels pro popisky), 01M je vybrané měřítko, 2016 je rok regulace statistické jednotky (aktualizace 14. 3. 2019) a 4326 značí EPSG:4326, také známé jako geografická projekce WGS84 v desetinách stupňů. LEVL odpovídá NUTS.

Ostatní soubory jsou v dalších geografických projekcích (EPSG:3035 pro Evropskou projekci LAEA a EPSG:3857 pro WGS84 / Pseudo-Mercator), menších územních celcích a rozlišené dle regionů, hranic a popisků.

Tento soubor kompletně rozbalíme:

```
$ unzip NUTS_RG_01M_2016_4326.shp.zip
```

Získáme tak hlavní soubor vektorových dat (shp), indexový soubor (shx) a databázový soubor (dbf). Tyto soubory jsou povinné, vedle toho získáme dva doplňkové soubory: zaznačení kódové stránky pro dbf (cpg) a přesné informace o projekci (prj).

3. Softwarová příprava

Pokud nemáme, nainstalujeme si R a potřebné knihovny, v našem případě minimálně `libssl-dev`, bude jej potřebovat plotly u interaktivní online verze:

```
$ sudo apt install r-base-core
$ sudo apt install libssl-dev
```

Ve výpočetním prostředí R si doinstalujeme knihovny:

```
$ R
> install.packages(c("sf","dplyr","ggplot2","ggrepel","plotly","tcltk"))
> q()
```

Ideální je instalovat knihovny jednu za druhou, skrz možné chybové zprávy.

Dále si vytvoříme soubor `robust-celek.R`, který si postupně okomentujeme.

Může se vám hodit TEXLive, ImageMagick nebo GraphicsMagick, prohlížeč rastrových obrázků (např. ristretto či GIMP) a prohlížeč pdf souborů (např. Okular či XpdfReader). Necháváme na laskavém čtenáři, jak hluboce se chce do článku ponořit.

4. Stati(sti)cká mapa

Načteme si potřebné knihovny.

```
library(sf); library(dplyr); library(ggplot2)
library(ggrepel); library(tcltk); library(plotly)
```

Připravíme si data o konferenční řadě v proměnné `rob`(ust).

```
### Data o ROBUSTech ###
rob = data.frame(
  location =  c("Načetín","Kost","Slavonice","Adršpach","Plasy",
    "Liblice","Herbertov","Malenovice","Lednice","Radešín",
    "Nečtiny","Hejnice","Třešť","Lhota nad Rohanovem","Pribylina",
    "Králíky","Němčičky","Jetřichovice","Kurzovní","Rybník","Bardějov"),
  longitude = c(13.2634230,15.1350694,15.3195072,16.1025647,13.3838904,
    14.5853101,14.3358220,18.4153645,16.8023564,16.0841326,
    13.1450920,15.2190330,15.4885601,13.6832665,19.7801583,
    16.7519537,16.8280667,14.3837522,17.2096600,12.6724633,21.2713128),
  latitude = c(50.5489359,50.4902238,49.0149020,50.6203202,49.9340679,
    50.3237900,48.6200863,49.5678849,48.8082447,49.4666355,
    49.9541085,50.8502057,49.3102300,49.1399375,49.1334503,
    50.1007519,48.9331628,50.8655207,50.0785817,49.5053812,49.2926572),
  Termín = c("letní","zimní","letní","zimní","letní",
    "zimní","letní","zimní","letní","zimní",
    "letní","zimní","letní","zimní","letní",
    "zimní","letní","zimní","letní","zimní","2020"),
  roky = seq(1980,2020,by=2)
)
```

Načteme si mapu a vytvoříme si základ v proměnné `p`(lotly).

```
### Základ mapy ###
map = st_read("NUTS_RG_01M_2016_4326.shp", stringsAsFactors=FALSE)
CZ0 = map %>% filter(NUTS_ID %in% c("CZ","SK"))
p = ggplot(rob) +
  geom_sf(data = CZ0, show.legend = FALSE, color="black") +
  xlab("") + ylab("") +
  geom_point(data=rob,
    aes(x=longitude, y=latitude, fill=Termín),
    pch=21, size=4, alpha=I(0.7)) +
  scale_fill_manual(values=c("green","red","blue")) +
  theme_void() +
  theme(
    legend.position="bottom",
    legend.title = element_blank(),
    panel.grid.major = element_line(linetype="blank"),
    panel.grid.minor = element_line(linetype="blank")
  )
```

Náš poslední krok je přidat popisky.

```
### Přidání popisku, statická verze ###
pfinal = p + geom_label_repel(data=rob, aes(x=longitude, y=latitude,
  label=location), cex=3)
```

Nyní se můžeme podívat na výsledek:

```
### Zobrazení v prostředí R ###
pfinal
```

Pokud bychom rádi mapu viděli přes `Rscript`, můžeme užít `X11()` pro operační systém Linux, resp. `windows()` či `quartz()` pro Microsoft Windows a Mac OS X. Pozastavení zrealizujeme přes knihovnu `tcltk`.

```
### Zobrazení přes Rscript ###
X11() # windows() nebo quartz()
pfinal
prompt  <- "Mezerníkem se zavře graf"
capture <- tk_messageBox(message = prompt)
dev.off()
```

Dále si ukážeme uložení do pdf.

```
### Uložení do PDF, verze tisková ###
cairo_pdf("Robust.pdf", width=10, height=7)
pfinal
dev.off()
```

Pokud bychom si přáli ořezat ochrannou bílou zónu, lze to udělat přes příkaz

```
$ pdfcrop --hires --margins 1 Robust.pdf
```

Nástroj `pdfcrop` je součástí TEXové distribuce TEXLive. Nula se zpravidla nepoužívá kvůli přesahům Bézierových křivek v písmech.

Jiná možnost je uložit si mapu do rastrového obrázku a ten si zobrazit přes prohlížeč obrázků.

```
### Uložení do PNG, verze na web ###
png("Robust.png")
pfinal
dev.off()
browseURL("Robust.png")
```

Pokud chceme ořezat ochrannou bílou zónu i zde, lze i to:

```
$ gm convert Robust.png -flatten -fuzz 1% -trim +repage Robust-crop.png
$ #  convert Robust.png -flatten -fuzz 1% -trim +repage Robust-crop.png
```

5. Interaktivní mapa offline

Náš další krok je připravit si verzi pro internet. Inspiraci lze hledat na serveru `https://plot.ly/r/`.

Prvně zasáhneme do vzhledu mapy.

```
### Přidání popisků, interaktivní verze ###
p = ggplot(rob) +
  geom_sf(data = CZ0, show.legend = FALSE) +
  xlab("") + ylab("") +
  geom_point(data=rob,
    aes(x=longitude, y=latitude, fill=Termín, text=paste(roky,'<br>',location)),
    pch=21, size=4, alpha=I(0.7)) +
  scale_fill_manual(values=c("green","red","blue")) +
  theme_void() +
  theme(
    legend.position="bottom",
    panel.grid.major = element_line(linetype="blank"),
    panel.grid.minor = element_line(linetype="blank")
  )
p <- ggplotly(p, tooltip= c("text")) # c("location"))
p <- p  %>%
  layout(legend = list(
    orientation = "h", x = 0.4, y =0
  ))
```

Zobrazit si mapu v R v offline režimu je jednoduché.

```
### Zobrazení mapy bez nutnosti účtu na plot.ly ###
p
```

Na první dobrou se nám nepodařilo mapu zobrazit přes `Rscript`. Jeden ze zajímavých tipů na internetových fórech[1] bylo místo `p` použít jednu z možností:

```
p %>% print
print(p)
show(p)
```

To sice začalo fungovat, pokud jsme blok kódu volali v R přes příkaz `source` a v RStudiu, ale nikoliv přes `Rscript`.

Řešení jsme objevili zde[2]. Nemáme-li, doinstalujeme si webové prohlížeče:

```
$ sudo apt install chromium-browser
$ sudo apt install firefox
```

Návod pro `chromium-browser` funguje bezvadně.

---

[1] Namátkou zde: `https://stackoverflow.com/questions/26643852/ggplot-plots-in-scripts-do-not-display-in-rstudio`

[2] `https://stackoverflow.com/questions/44048347/r-open-plotly-in-standalone-window`

```
81 print_app <- function(widget) {
82   temp <- paste(tempfile('plotly'), 'html', sep = '.')
83   htmlwidgets::saveWidget(widget, temp, selfcontained = FALSE)
84   system(sprintf("chromium-browser -app=file://%s", temp))
85   # system(sprintf("firefox file://%s", temp))
86   # system(sprintf("sleep 5"))
87   temp
88 }
89 print_app(p)
```

Ovšem u prohlížeče `firefox` se dočasný soubor smaže dřív, než jej prohlížeč dokáže otevřít. Řešení je pozastavit funkci `print_app` a jsou to zakomentované řádky 85 a 86.

6. Interaktivní mapa online

Náš poslední robustní úkol je mapa v online režimu.

- Otevřeme si `https://chart-studio.plot.ly/`.
- Naklikneme Sign Up.
- Vyplníme potřebné registrační údaje. Dáme SIGN UP.
- Potvrdíme údaje přes odkaz v došlém emailu.
- Po najetí na vytvořené uživatelské jméno (vpravo nahoře) dáme Settings.
- Vlevo v menu zvolíme API Keys.
- Přes Generate Key si necháme vytvořit API klíč.
- Někam si uživatele a vytvořený API klíč uložme. Hned je použijeme. Heslo je dobré si též uložit, ale v R jej potřebovat nebudeme.

Vrátíme se zpět do zdrojového kódu a zapišme:

```
### Zobrazení mapy na serveru plot.ly ###
Sys.setenv("plotly_username"="VÁŠ UŽIVATEL") # <-- zasáhněte
Sys.setenv("plotly_api_key" ="VÁŠ API KLÍČ") # <-- zasáhněte
chart_link = api_create(p, filename="Robust-2020")
chart_link
```

Proběhlo-li vše v pořádku, otevře se webový prohlížeč a vznikly (nebo se aktualizovaly) dvě záložky: Robust-2020 Grid a Robust-2020.

Pracovní verze mapy je dostupná přes stránku `https://chart-studio.plot.ly/create/?fid=vencalek:5#/`, zkrácený odkaz je `plot.ly/~vencalek/5`.

Nezávislý test je uložen na `https://chart-studio.plot.ly/create/?fid=malipivo:3#/`, resp. `plot.ly/~malipivo/3`.

Gratulujeme! Jsme hotoví.

7. Mapa na produkčním serveru

Webové stránky ČStS – statspol.cz – jsou přepracované. Pokud z menu vybereme KONFERENCE a dále ROBUST, mapa se otevře v pravé horní části. Přímý odkaz na výsledek našich snah je http://www.statspol.cz/konference/robust/.



Reference

Knih a materiálů je nespočet, zájemce o grafiku v R s knihovnou plotly odkazujeme na nový knižní přírůstek z ledna 2020 [1].

[1] Sievert, C.: *Interactive Web-Based Data Visualization With R, plotly, and shiny*. 1st edition. Chapman & Hall/CRC The R Series, January 21, 2020, pp. 448. On-line version available at: https://plotly-r.com/. ISBN 978-1138331457.

Kontaktní adresy

doc. Mgr. Ondřej Vencálek, Ph.D., Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzity Palackého, 17. listopadu 12, 771 46 Olomouc,
E-mailová adresa: ondrej.vencalek@upol.cz

Ing. Pavel Stříž, Ph.D., U Škol 940, Bučovice, okres Vyškov, 685 01, Česká republika,
E-mailová adresa: pavel@striz.cz